APPENDIX XIII
III

LIMITATIONS OF IBM ACCOUNTING MACHINES

Searching for patterns as conducted by scanning machines stands in sharp contrast with sorting operations performed by conventional punched-card machines designed for accounting. When using a conventional sorter to locate a desired entry, its position on the card must be known or determined. In other words, for conventional accounting purposes, meaning is ascribed to punched holes by designating their geometrical coordinates, namely, column number (the longitudinal dimension) and position in the column (the latitudinal dimension).

When using IBM sorters and collators of conventional design, decisions must be made in advance as to which column (or sequence of columns) will be used for recording a given type of data such as, for example, the number of units of a given item sold, its price, the date shipped, the discount if any, the total amount owed by the customer, the code number of the salesman, the point from which shipped. In addition to assignment of position the code for a given entry must be fixed as to the number of symbols of which it consists. By meeting these requirements as to location of a given entry and number of symbols involved it becomes possible for conventional sorting equipment to distinguish between different meanings attributed to the same punching in different locations. Thus, for example, the numeral 5 punched in different predetermined columns may be attributed such meanings as a unit price is $5.00 per item or the number of items shipped or salesman No. 5 responsible for the Rocky Mountain Regions, etc. In accounting procedures, specific types of data such as unit price, salesman code number are predominate and are known in advance. It is therefore possible to assign certain columns or sequences of columns to specific types of data in general. This approach suffices for accounting purposes. A shipment of some one item will involve some one amount of the item sold at some one price going to a given customer located at some one place. As a consequence, preassignment of fixed areas for specific data entries is able to provide for many of the requirements of accounting.

The establishment of fixed locations for punched data, although acceptable in accounting, imposes limitations that must be taken into account when adapting conventional IBM equipment to information searching. If the information to be searched is restricted to some one narrow field of specialization, then it is possible to reserve a location on the card for each entry, or group of mutually exclusive entries, that might be required to designate the subject matter of the documents. As the field to be

CONFIDENTIAL
Security Information

covered becomes broader, the total number of entries required
becomes greater and it becomes impossible for conventional
punching methods to provide on a standard IBM card enough
locations to accommodate the whole range of possible entries.
This means that standard IBM equipment would require an ex-
cessive number of cards to provide enough locations to accommo-
date a wide range of documents. As a consequence, conventional
punching employing fixed fields cannot be used if index entries
pertaining to a single document are to be recorded as a block
so as to facilitate searching to new combinations of entries
corresponding to new points of view or to unanticipated require-
ments.

The alternative--as worked out in applying IBM accounting
equipment to Agency requirements--has been to enter only one
subject entry on an individual card. This eliminates the
possibility of conducting searches in such a way that machine
selecting operations are directed to a new combination of
concepts, that is to say to a combination of concepts not
previously set up as a complex subject heading. As a con-
sequence, in the Agency's present machine system, it is
possible to define and to direct searching operations only in
terms of such combinations of headings as may have been established
prior to analysis of information preparatory to punching the cards.
In other words, new combinations of subject headings cannot,
in any practical manner, be made the basis of machine searching
operations. Experience has demonstrated that these restrictions
have imposed intolerable limitations on the ability of the
present OCD system to meet the continually shifting needs of the
analysts of OSI and ORR.

As already noted, scanning machines conduct a search by
detection of specific patterns which correspond to those index
entries which are used to define the scope of the search.

For maximum flexibility, any machine searching system must
be able to express the subject matter of each document by an
appropriate array of index entries. This means, from a practical
point of view, that no restrictions may be imposed on the combina-
tions of entries assigned to a given document. With scanning
machines, such restrictions are, in fact, avoided as any index
entry may be spelled out by appropriate patterns and recorded on
the searching medium without restriction as to locations. There
is no need, as in the case with standard tabulating equipment,
to reserve for a given entry or class of entries, a certain
location--or field, as it is sometimes called.

CONFIDENTIAL
Security Information

Machine Indexing Project

Interim Report

est date August '53

004

24/3/78   022